Draft: 6/24/21

**Data Survey Definitions**

**DRAFT**

**Artificial Intelligence/Machine Learning (AI/ML)**

AI/ML describes an automated process in which a system begins recognizing patterns without being specifically programmed to achieve a pre-determined result.  This is different from a standard algorithm in that an algorithm is a process or set of rules executed to solve an equation or problem in a pre-determined fashion.  Evolving algorithms are considered a subset of AI/ML.

**Artificial Intelligence / Machine Learning Systems include:**

* Systems that adapt and adjust to new data and experience without manual human intervention.
* Systems that arrive at results for which the outcomes and the stepwise approach toward the outcomes were not configured in advance by a human programmer.
* Systems that dynamically respond to conditions in the external environment without the specific nature of such responses being known in advance to the designers of the systems.
* Systems that utilize neural networks and/or deep-learning algorithms, such as supervised, semi-supervised, and unsupervised learning algorithms.
* Systems that engage in automatic speech recognition, facial recognition, image recognition, text recognition, natural language processing, generation of customer-specific recommendations, automated customer communications (e.g., chatbots with non-preprogrammed prompts), autonomous or semi-autonomous vehicle operation or data gathering, or any other approach that does not require either preprogramming or a manual human intervention in every instance of an action or decision.
* Systems that automatically generate adaptive responses based on interactions with a consumer or third party.
* Systems that determine which data elements to rely upon, in a non-preprogrammed fashion, among a variety of possible alternatives.

**Artificial Intelligence / Machine Learning Systems are not:**

* Static “scorecards” that deterministically map consumer or other risk characteristics to treatments or decisions. (However, an AI/ML system may use the output of such static “scorecards” as input data for the AI/ML system to consider.)
* Systems with solely preprogrammed decision rules (e.g., “If A, then B” applied invariably in all situations).
* Tables of point or factor assignments in rating plans.
* Static ratemaking and/or predictive-modeling methodologies, including linear regression, generalized linear modeling (GLM), or generalized additive modeling (GAM).
Purely informational static databases, such as databases used to obtain reference amounts for claim settlements, or static databases pertaining to consumer characteristics or experience, regardless of the amount of information in the database.  However, if AI/ML is used to create a static predictive model, that AI/ML system is considered within the scope of this survey.
* Deterministic “phone trees” that navigate consumers through pre-recorded voice prompts.
* Any approach that an insurer could have realistically utilized in the year 2000 or prior.

**AI/ML Use Descriptions and/or Explanations**

* **Rating:** AI/ML Uses
	+ Rating Class Determination:
	+ Price Optimization: As defined in the NAIC Casualty and Actuarial Statistical (C) Task Force white paper: <https://www.naic.org/documents/committees_c_catf_related_price_optimization_white_paper.pdf>
	+ Retention Modeling:
	+ Numerical Relativity Determination:
* **Underwriting:** AI/ML Uses
	+ Automated Approval:
	+ Automated Denial:
	+ Underwriting Tier Determination:
	+ Company Placement:
	+ Input into Non-Automated Approval Decision:
	+ Input into Non-Automated Denial Decision:
	+ Automate Processing Thru the Agency Channel:
* **Claims:** AI/ML Uses
	+ Claim Approval:
	+ Claim Denial:
	+ Determine Settlement Amount:
	+ Claim Assignment Decisions:
	+ Informational Resource for Adjusters:
	+ Evaluation of Images of the Loss:
* **Fraud Detection:** AI/ML Uses
	+ Fast Tracking of Likely Non-Fraudulent Claims:
	+ Referral of Claims for Further Investigation:
	+ Detect Medical Provider Fraud:
	+ Detect First-Party Liability:
	+ Detect Third-Party Liability:
* **Marketing:** AI/ML Uses
	+ Targeted Online Advertising:
	+ Identification of Recipients of Mail or Phone Advertising:
	+ Provision of Offers to Existing Customers:
	+ Identification of Potential Customer Groups:
	+ Demand Modeling:
	+ Direct Online Sales:
* **Loss Prevention:** AI/ML Uses: With advances in technology, the loss-prevention function is likely to grow in importance as insurers strive to improve their results not only through efforts at risk selection and matching the rate to the risk, but also through reducing the probability that the underlying insured losses would occur in the first place. Possible uses of artificial intelligence in loss prevention for private passenger automobile insurance may include, but are not limited to, the following:
	+ Identification of High-Risk Customers: The goal of such identification in a loss-prevention context is not to make an underwriting or rating decision, but rather to recognize which specific customers may benefit most from loss-prevention advice and mitigation techniques that the insurer may be able to provide, thereby reducing such customers’ frequency and/or severity of losses. For example, an AI/ML system might determine that certain households with youthful drivers are more likely to benefit from risk-mitigation advice and other approaches.
	+ Risk-Mitigation Advice to Consumers: Artificial intelligence systems might be used to target messaging to consumers based on specific risks identified for a given policy. For example, in a household with youthful drivers, AI/ML-targeted messaging and incentives could focus on ways those drivers could gain experience in a low-risk manner and drive more carefully in day-to-day context. For households in mountainous areas, AI/ML systems could provide targeted advice about safe driving in rugged terrain.
	+ Determination of Advance Payments: In many situations, small payments issued at or shortly after the time of loss, prior to the full adjustment of the claim, can help the insured or third-party claimant prevent much larger amounts of damage that would otherwise greatly raise the costs of the claim for the insurer. In a private passenger automobile context, examples could include, but are not limited to, (i) making a payment for minor repairs that restore the vehicle to a drivable condition, whereas the insured and/or insurer would have otherwise needed to spend much more money to rent another vehicle or to pay for storage of a non-functional vehicle; (ii) making a payment for prompt, inexpensive medical treatment of a claimant, which could prevent the emergence of a longer-term, chronic, and much more costly health condition; or (iii) making a payment for expenses related to towing an insured’s or claimant’s vehicle away from the scene of the accident and reasonable costs of storage for the vehicle until the insurer or vehicle owner are able to gain possession of the vehicle. In the absence of such prompt payments, vehicles at towing-company storage yards may accumulate significant charges for which the insurer may ultimately become responsible.

**Data Use Definitions**

1. Criminal Convictions (exclude auto-related convictions)
2. Demographic (age, gender, address, marital status, other non-behavioral attributes of a consumer or population attributes of an area)
3. Driving Behavior (tickets, years of driving experience, annual miles driven)
4. Education (level of education, GPA)
5. Vehicle-Specific Data (type of vehicle(s) driven or owned, history of the vehicle(s), value of contents inside the car)
6. Facial Detection / Recognition / Analysis (picture to confirm identity, estimate biological age or gender of the consumer)
7. Geocoding (latitude and longitude coordinates of a physical address)
8. Natural Catastrophe Hazard (frequency and severity of natural hazards)
9. Job Stability (current employment, length of employment at prior employers, unemployment)
10. Income (annual income, income source)
11. Occupation
12. Personal Financial Information (net worth, type of bank account or credit account, number of bank accounts or credit accounts, available credit, payment history data)
13. Loss Experience (claim history for PPA, claims from other lines of insurance)
14. Medical (medical history, medical condition, prescription data, lab data)
15. Online Media (web searches, online purchases, social media activities)
16. Telematics (time-of-day driving data, location of driving data, braking data, acceleration data, maximum spend, turn speed)
17. Voice Analysis (speed, pitch, volume)]
18. Consumer or Other Type of “Score”: A numeric value generated based on a combination of any underlying attributes or behaviors of the consumer, insured risk, or any items considered by the insurer to be relevant to the consumer or insured risk. Scores are computed using deterministic algorithms or models which are not themselves considered to be AI / ML systems. Inquiries in this survey regarding such scores seek to understand whether these scores are used as input data elements within AI / ML systems.
19. Other

**Data Use Tab Example**

Example:

* + **ROW 3:** Rating Section: Put a “Y” in the box for every category of data used in an AI/ML system for a Rating use.
	+ **ROW 3:** For columns S and T, put a “Y” in the box IF a consumer score for a category not specified in columns B – R is being used and put a “Y” in the box if some type of “Other: Non-Traditional Data Elements” are being used not specified in columns B – R.
	+ **ROW 4:** Put a Y in the box for every category of data used in an AI/ML system for Rating use that is a derived “score” representing that data category (for example, a consumer “education” score or a driving behavior “telematics” score.
	+ **ROW 5:** For each category of data element used in an AI/ML system (where there is a “Y” in the box), indicate whether it is sourced internally (I), externally (E) or both internally and externally (B). This pertains to data elements, not “scores”.
	+ **ROW 6:** If Row 5 indicates “E” for externally sources or “B” for sources internally and externally, list the vendor.
	+ **ROW 6:** In column S, if in Row 5 for column S there is either an “E” or “B”, indicating consumer or other “scores” used are sourced externally or both internally and externally, list the score data providers.

**Other AI/ML Use Descriptions:**

* What is the difference between a Proof of Concept (POC) and a Prototype?

A Proof of Concept (POC) is a small exercise to test the design idea or assumption. The main purpose of developing a POC is to demonstrate the functionality and to verify a certain concept or theory that can be achieved in development. Prototyping is a valuable exercise that allows the innovator to visualize how the product will function. A prototype is a working interactive model of the end-product that gives an idea of the design, navigation, and layout.  While a POC shows that a product or feature can be developed, a prototype shows how it will be developed.